

H.264 동영상 압축의 문법적 제한요소를 이용한 MAP기반의 Channel Decoder 성능 향상에 대한 연구

A Study on Channel Decoder MAP Estimation Based on H.264 Syntax Rule

전 용 진*, 서 동 완*, 최 윤 식**

* 연세대학교 전기전자 공학과(전화:(02)2123-2774, 팩스:(02)313-0960, E-mail : lorenzo@yonsei.ac.kr)

* 연세대학교 전기전자 공학과(전화:(02)2123-2774, 팩스:(02)313-0960, E-mail : md9531@hanmail.net)

**연세대학교 전기전자 공학과(전화:(02)2123-2774, 팩스:(02)313-0960, E-mail : yschoe@yonsei.ac.kr)

Abstract : In this paper, a novel maximum a posteriori (MAP) estimation for the channel decoding of H.264 codes in the presence of transmission error is presented.

Arithmetic codes with a forbidden symbol and trellis search techniques are employed in order to estimate the best transmitted.

And, there has been growing interest of communication, the research about transmission of exact data is increasing. Unlike the case of voice transmission, noise has a fatal effect on the image transmission. The reason is that video coding standards have used the variable length coding. So, only one bit error affects the all video data compressed before resynchronization.

For reasons of that, channel needs the channel codec, which is robust to channel error. But, usual channel decoder corrects the error only by channel error probability.

So, designing source codec and channel codec, instead of separating them, it is tried to combine them jointly. And many researches used the information of source redundancy in received data.

But, these methods do not match to the video coding standards, because video coding standards use not only one symbol but also many symbols in same data sequence.

In this thesis, We try to design combined source-channel codec that is compatible with video coding standards. This MAP decoder is proposed by adding semantic structure and semantic constraint of video coding standards to the method using redundancy of the MAP decoders proposed previously. Then, We get the better performance than usual channel coder's.

Keywords : MAP decoder, Joint Source-Channel Decoder

1. 서 론

높은 압축률을 가지는 동영상 압축 방식의 필요성에 의하여 2001년 MPEG에서는 VCEG와 공동으로 JVT(Joint Video Team)를 결성하였다.

JVT의 주 목적은 새로운 국제 표준안인 H.264/AVC (Advanced Video Coding)을 제정하는 것이었다. 현재 제정된 H.264는 기존의 동영상 압축 표준안보다 압축 성능이 훨씬 우수한 것으로 알려져 있다[5].

동영상 전송에 있어서 채널의 오류는 치명적인 영향을 미칠 수가 있다. 고정부호화기(Fixed Length Coding, FLC)의 경우, 채널의 오류에 의한 영향은 단지 하나의 코드워드에만 국한되지만, 요즘의 영상압축 표준은 가변부호화기(Variable Length Coding, VLC)를

기본으로 하기 때문에 단 한 비트의 채널오류에 의한 영향이 다음 재동기화(Re-synchronization)가 일어나기 전까지 전파되어 안정적인 영상을 얻기가 힘들어진다.

소스 부호화기는 소스의 중복성을 최대한 줄여서 전송하고, 채널 부호화기는 채널오류에 의한 영향을 줄이기 위해 부가정보를 추가함으로써 이루어진다[1]. 이는 Shannon에 의한 separation theorem에 의한 것으로, 채널 부호화기는 모든 오류를 정정할 수 있다는 가정과 부호화기 설계에 있어서의 복잡성에 어떠한 제한도 존재하지 않는다는 조건하에 소스 부호화와 채널 부호화는 최적화에 지장 없이 분리되어 부호화가 가능하다. 하지만, 실제 시스템의 경우 이러한 분리는 가능하지 않다[2].

이런 이유로 채널과 소스의 관계를 분리시키지 않고 서로 연관시켜서 더 나은 성능을 얻으려는 연구들이 이루어지고 있다[3].

본 논문에서는 채널의 오류 확률과 소스의 전이 확률 분포를 이용하여 채널 잡음에 의해 손상된 정보를 올바르게 정정하는 방법으로 동영상 압축 표준들이 가지는 문법적인 제한과 문법적인 구조를 이용해 H.264의 동영상압축표준에 적용 가능한 복호기를 설계하였다[5][6][7].

II. H.264의 소스 채널 부호화 및 부가정보

1. 소스 채널결합 복호화기

크기 N 의 알파벳 $J_N = \{0, 1, \dots, N-1\}$ 을

가지는 stationary stochastic process $\{J_n\}_{n=1}^{\infty}$ 이 모든 $n > 1$ 에 대해 다음 조건을 만족하면, discrete Markov source라고 말한다.

Sequence MAP detection

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \Pr\{J_1^n = i_1^n | J_1^n = j_1^n\} \quad \text{Most Probable Transmitted Sequence}$$

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \frac{\Pr\{J_1^n = j_1^n | J_1^n = i_1^n\} \Pr\{J_1^n = i_1^n\}}{\Pr\{J_1^n = j_1^n\}} \quad \text{Bayes' Theorem}$$

$$\hat{i}_1^n = \arg \min_{i_1^n \in J_N^n} \left[\prod_{k=1}^n Q(i_k | i_k) \right] \left[\prod_{k=2}^n P(i_k | i_{k-1}) P(i_1) \right] \quad \text{Markovian Property}$$

$$\hat{i}_1^n = \arg \max_{i_1^n \in J_N^n} \left[\sum_{k=2}^n \log[Q(i_k | i_k) P(i_k | i_{k-1})] + \log[Q(i_1 | i_1) P(i_1)] \right]$$

convolutional 복호화기의 path metric과 비슷한 형태이다. 위의 식은 branch metric이 두 가지로 구성되는데 이는 $\log[Q(i_k | i_k)]$ 과 $\log[P(i_k | i_{k-1})]$ 이다.

$\log[Q(i_k | i_k)]$ 항은 채널에 전적으로 의존하는 식으로 i_k 를 보냈을 때, i_k 를 받을 확률을 나타내고, $\log[P(i_k | i_{k-1})]$ 항은 소스의 확률 분포에 의존한다. 즉, $k-1$ 번째에 i_{k-1} 을 보냈을 때, k 번째 i_k 를 보내는 확률을 나타낸다[2][3].

2. H.264에서의 부호화 방식 및 여러 전파 모델

H.264에서는 보다 효율적인 예측 부호화를 위하여 16X16 매크로 블록을 일정한 크기와 모양으로 최소 4X4 부블록까지 분할하여 부블록 내의 모든 화소들에 대하여 하나의 움직임 벡터를 구하고 움직임 보상 매크로 블록 모드는 부호화할 매크로 블록당 최대 16개의 움직임 벡터 정보가 추가적으로 요구된다.

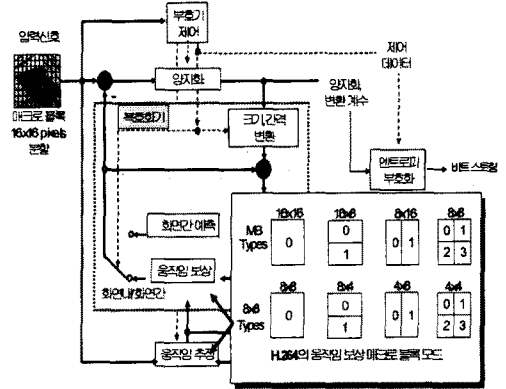


그림1. H.264에서의 움직임 보상 블록도

또한 화면 내 예측 모드와 더불어 움직임 보상 매크로 블록 모드 정보도 부가적으로 요구된다.

동영상압축과정에서 흔히 쓰이는 DCT나 움직임 보상 알고리즘에는 가변 부호화를 사용한다. 따라서 일단 이러한 가변 부호에 발생한 오류는 복호과정에서 이후 데이터에까지도 영향을 미쳐서 복호화기의 재동기가 이루어지기 전까지 다른 블록들에 영향을 미치고 이렇게 공간적으로 전파된 오류는 움직임 보상등에 의해 다음 픽처로 전파된다.

H.264에서는 매크로 블록들을 그림의 한 열 단위로 모아 슬라이스라는 단위를 두고 있는데 슬라이스 계층 내의 데이터 열은 대부분 가변 부호로 이루어져 있고 복호화기의 재동기는 픽처 개시부 혹은 슬라이스 개시 부호에 의하여 이루어진다[5]. 따라서 슬라이스헤더보다 하위계층에서 오류가 발생하면 그 영향이 매크로 블록 또는 슬라이스내에서 전파가 이루어 질 수 있다.

III. H.264에서의 MAP기반의 소스 채널 결합 복호화기

동영상 압축표준은 여러 개의 계층으로 이루어져 있어 그림 한 장 또는 시퀀스의 정보를 가지고 있는 헤더와 그 그림들의 픽셀값을 표시하는 데이터부분으로 나누어진다. 데이터부분에 오류가 발생하면 오류의 영향은 적게는 그 데이터가 포함된 매크로 블록에만 한정될 수 있지만, 헤더부분에 오류가 발생할 경우 그 오류의 영향은 그림 전체에 영향을 미친다.

오류의 전파를 최소화 하기위해 그림2와 같이 제안하는 과정을 통해 소스의 심볼이 갖는 문법적 제약요소와 확률모델을 통해 오류의 전파를 줄여 복호화기의 성능을 향상 시키고자한다.

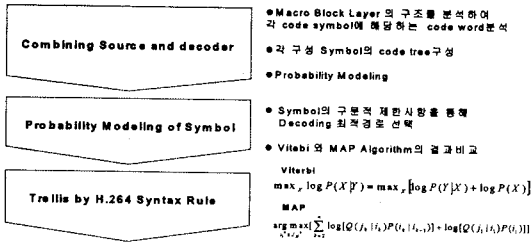


그림2. 소스 채널 결합 복화기의 제안 알고리즘

1. 가변부호의 소스 채널 복화화

부분의 데이터 정보는 가변부호를 사용하는데 실험을 통해서 가장 많이 나오는 코드에 대해서는 적은 비트를 할당하고 적게 나오는 코드에 대해서는 많은 비트를 할당하는 허프만 코드를 사용하여 만들어진 것이다.

'00', '01', '10', '11'의 코드를 갖는 그림 3과 같은 허프만 코드 트리의 경우 네 개의 출력 중 하나가 나올 확률은 모두 1 / 4로 같고, 어떤 상태에서 하나의 입력이 들어와 다른 상태로 옮겨갈 확률은 각각 1 / 2로 같다. 하지만 대부분의 심볼이 가변부호이므로 하나의 출력이 나올 확률은 같지 않다.

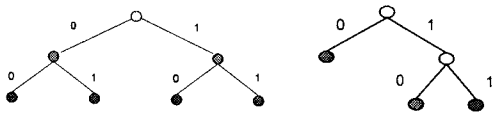


그림3. 허프만 코드트리 구조 (좌도: 고정부호, 우도: 가변부호)

우도의 경우 '0'이 나올 확률은 1 / 2, '10'과 '11'이 나올 확률은 1 / 4이다. 또한 한 상태에서 다른 상태로 옮겨갈 확률 즉, '1'이 나온 후 '0'이나 '1'이 나올 확률은 1 / 2이다. 이처럼 모든 헤더 정보들과 데이터 정보들은 각각의 정보에 맞게 주어진 허프만 코드를 사용한다. 기존의 논문은 움직임 벡터의 잉여 정보를 이용하였는데, 이때 움직임 벡터들은 가변부호로 부호화되었고, 연습(training)을 통해 하나의 소스 출력에서 다른 소스 출력으로 옮겨갈 수 있는 확률을 정의하였다[3].

2. H.264의 문법적 구조 및 제한요소 모델

문법적인 구조는 한 심볼에서 다른 심볼로 옮겨갈 때 주어지는 제한을 말한다. 기존의 연구에서처럼 움직임 벡터만의 심볼이 연속적으로 오는 것이 아니고, 그림 4와 같은 구조로 심볼이 들어오게 되는데 문법적

인 연속성을 다음의 경우들로 그 가능성을 한정시켰다.

Viterbi 복화화기가 한 채널 상태로 들어오는 모든 경우의 가능성에서 가장 큰 확률을 가지는 경우만 선택하듯이 MAP 복화화기도 Viterbi 복화화기의 경우와 같이 한 심볼의 첫 번째 비트로 들어오는 경로들 가운데 가장 큰 가능성을 가지는 경로만을 선택한다. 이와 함께 아래의 H.264의 슬라이스 층의 주요한 문법적인 제한조건 예로부터 제한 조건들을 이용하면 복화화기에서 비트 스트림을 Viterbi 복화화기에서 일어날 수 있는 오류에 의한 문법적인 오류를 막을 수 있는 장점이 있다.

- * Ref Pic ID 이후 MB이 연속해야한다
- * RPSL이 '1'이면 MB Layer Read가 수행되지 말아야 한다.
- * MC Process에서 RMPNI index가 '0'이면 ADPN Loop를 수행해야한다

이러한 구조와 함께 심볼내의 소스 전이 확률은 각각의 심볼마다 주어진 허프만 코드의 길이에 의한 확률로 정의하였다[5][6].

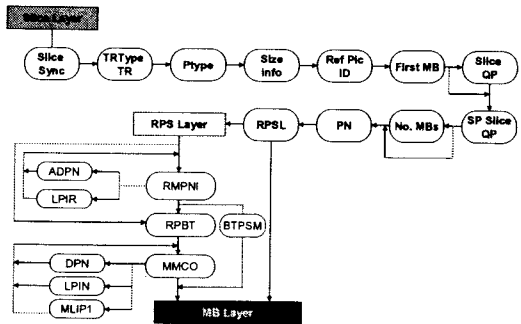


그림4. H.264의 슬라이스 층의 문법적 구조

IV. 실험 및 고찰

본 논문에서 제안한 MAP 복화화 방법과 기존의 Viterbi 복화화 방법에 의한 성능을 비교하기 위해 그림 5와 같은 실험을 수행하였다. 오류에 의한 영향으로 인해 그림 전체가 영향을 받아 정확한 성능 비교를 할 수 없는 상황을 피하기 위해 오류에 의한 영향을 최소화한 옵션을 사용하였다. 실험 영상은 Foreman과 Susie를 사용하였고, I-픽처에 오류가 발생할 경우 I-픽처뿐 아니라 P-픽처에도 오류가 발생하게 되는데 여기서는 심볼을 정의함으로써 얼마나 많은 성능향상을 볼 수 있는가

를 보고자하므로 P-픽처가 참조하는 I-픽처에는 오류를 삽입하지 않았으며, P-픽처에만 오류를 삽입하였다. 오류에 의한 직접적인 영향을 비교하기 위해 오류를 삽입한 첫 번째 P-픽처만을 고려하여 성능 향상 여부를 비교하였다. 또한, 헤더에 오류가 발생할 경우 오류에 영향을 받은 영상은 복호화가 불가능하므로 헤더정보는 더 강하게 채널 부호화가 되었거나 base-line이 아닌 enhancement-line을 통해 전송된다고 가정한다.

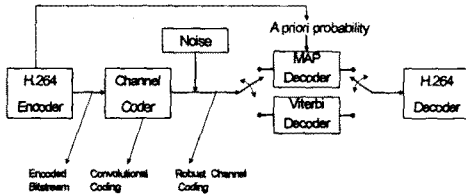


그림5. H.264의 소스 채널 결합 실험 모델

다음의 표1 에서 나타나는 PSNR를 보면, 영상에 따라 어느 정도의 차이는 있지만, Viterbi 복호화기에 의한 오류 정정에 비해 문법적인 제한을 가지는 MAP 복호화기가 더 나은 성능을 나타냄을 알 수 있다.

또한, 그림 6으로부터 MAP 복호화기가 좋은 성능을 나타냄을 확인할 수 있으며, 오류율이 증가할수록 Viterbi 복호화기와 비슷한 결과를 나타냄을 알 수 있다.

영상:Foreman			
H.264 [dB]	오류율	Viterbi 복호화기[dB]	MAP 복호화기[dB]
36.44	0.00367	32.53	33.62
36.44	0.00642	31.92	33.28
36.44	0.00989	31.05	32.34
36.44	0.0121	30.03	31.45
36.44	0.0278	28.23	29.33
36.44	0.0576	23.78	25.27

표 1. Foreman결과 영상의 PSNR

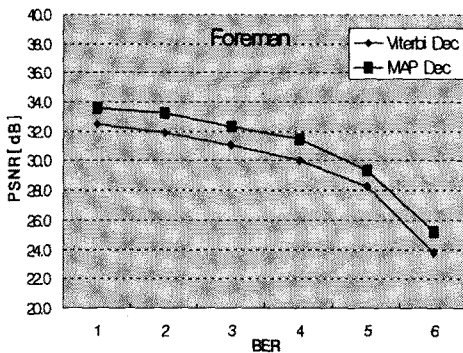


그림6. Foreman 결과영상의 PSNR

IV. 결 론

본 논문에서는 H.264과 같은 동영상 압축 표준을 오류가 존재하는 채널을 통과시킬 때, 채널오류로부터 영상 데이터를 최대한 정확하게 복원해 내는 방법을 제시하였는데, 이는 각 압축 표준에 존재하는 문법적인 구조를 이용하여 심볼들 사이의 가능한 전이 경로를 한정하였다. 또한, 문법적인 구조와 함께 문법적인 제한조건을 두어 문법적인 구조를 만족하는 경로들에서도 제한조건을 만족하는 최적의 경로를 선택하였다.

그 결과 Viterbi 복호화기에 의해 발생할 수 있는 문제인 동영상 압축 표준에 존재하지 않는 코드를 소스 복호화기에 보내 소스를 복호화하지 못하는 현상을 막을 수 있었다. 하지만, 높은 오류율의 경우 이러한 문법적인 제한에 의한 MAP 복호화기도 정확한 복호화는 불가능함을 확인하였다. 이러한 결과는 제안하는 MAP 복호화기 방법에서 소스내의 전이 확률을 단순히 심볼의 길이에 의해 부여했다는데 돌 수 있다.

또한 본 논문에서는 단순히 심볼의 길이에 의해 소스 전이 확률을 정의하였으나 연속적인 연습을 통해 더욱 정확한 소스 전이 확률을 갱신해 나가는 방법을 이용하는 연구들이 이루어져야 할 것이다.

참 고 문 헌

- [1] C. E. Shannon, "A mathematical theory of communication", Bell Syst. Tech. J., vol. 27, 1948.
- [2] A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding, 1979
- [3] K. Sayood and J. C. Borkenhagen, "Use of Residual Redundancy in the Design of Joint Source/Channel Coders", IEEE Trans. Comm., vol. 39, No. 6, June 1991.
- [4] Demir and K. Sayood, "Joint Source/Channel Coding for Variable Length Codes", DCC '98 Proceedings, 1998.
- [5] Thomas Wiegand, Gary Sullivan, Ajay Luthra "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)" Geneva May 27, 2003
- [6] Detlev Marpe, Heiko Schwarz, Blattermann, Guido Heising, and Thomas Wiegand "Context Based Adaptive Arithmetic Coding In JVT/H.26L" IEEE 2002
- [7] Detlev Marpe, Heiko Schwarz, and Thomas Wiegand "Context Based Adaptive Binary Arithmetic Coding In H.264/AVC Video Compression Standard" Manuscript May 21, 2003